ANALYSIS

29(106), March 1, 2015

ISSN 2278–5469 EISSN 2278–5450

Discovery

Design and Development of on Paper Digital Signature (OPDS) using Content Image

Sajan Ambadiyil[1]☼, Varun Rajeev[2], AvinashVakkalagadda[3], Pradyumna Kummamuru[3], Mahadevan Pillai VP[4]

1.Center for Development of Imaging Technology, Thiruvanthapuram-695027, Kerala, India
2.BITS Pilani, Pilani Campus, Pilani – 333031, Rajasthan, India.
3.BITS Pilani, Hyderabad Campus, Hyderabad - 500078, Andhra Pradesh, India.
4.Department of Optoelectronics, University of Kerala, Thiruvanthapuram-69558, Kerala, India

☼**Corresponding author:** Center for Development of Imaging Technology, Thiruvanthapuram-695027, Kerala, India; Email: ambadycdit@gmail.com

Publication History
Received: 04 January 2015
Accepted: 09 February 2015
Published: 1 March 2015

Citation
Sajan Ambadiyil, Varun Rajeev, AvinashVakkalagadda, Pradyumna Kummamuru, Mahadevan Pillai VP. Design and Development of on Paper Digital Signature (OPDS) using Content Image. *Discovery*, 2015, 29(106), 2-6

ABSTRACT

Valued paper documents are vulnerable to content manipulation in spite of several existing high-end security features. They are making way for the more advanced electronic documents. However, no matter the sophistication of paperless office they cannot efficiently replace paper documents. The on-paper digital signature scheme is a means for authenticating the content of a message and the signer's identity. The currently existing method employs the use of OCR technology while generating and verifying a digital signature. However, for higher accuracy, the method demands high-end OCR technology, which is cost prohibitive. The paper presents a novel cost effective method to extend the high security features of digital signature on to paper by considering the data as an image and converting to binary form. This method maintains all the security features of digital signatures like integrity, non-repudiation and authentication.

Keywords– Authentication, Digital Signatures, Hash function, QR Code.

1. INTRODUCTION

With advanced scanning and printing technologies, paper document fraud is becoming more prevalent. The various existing security measures include usage of special paper substrates and inks, unique printing methods, holograms, visual cryptography [1], watermarks, security fibers etc.[2].These methods tackle some fraudulent activities but they do not provide for content integrity. Methods that provide for tampering, forgery or counterfeiting include geometric lathe work like guilloche, micro printing, color changing chemical and inks on paper, magnetic inks etc. [3].For example, if a person tries to alter a certificate by using chemicals like acetone or kerosene on a certificate that has been made using the above mentioned special materials, then there will be a color change or a distinguishable mark on the certificate. But due to the cost prohibitive nature involved for the production of such special security document and effectiveness of implementing such system in to the confidential document, it is not possible to implement these techniques in all the documents. Apart from these constraints, the above mentioned methods also do not provide for content integrity. For example, if a person has access to such blank security paper, he can create his own document. On-paper digital signatures tackle the above concern.

Digital signatures are different from electronic signatures, since the latter do not depend on the content of the message. For a specific person, the electronic signature remains the same irrespective of the content. So, an electronic signature can be copied from one document to another. Digital signatures depend not only on the signer but also on the content of the message. So, a digital signature cannot be copied from one document to another [4].

A digital signature is basically a mathematical scheme that can ensure the integrity of the content of a message and also authenticates the sender's identity. It is a string of bits. They are implemented on paper in the form of QR codes. These QR codes are easily generated, printed on paper and can be easily verified using QR code verifier or even using mobile phones with a camera. The errors are detected and corrected using Reed-Solomon error correction.

Some practices implement digital signatures using an OCR to read the characters efficiently and then process them to create a digital signature. This creates the problem of efficiently recognizing the characters. An OCR is used to read the characters from the paper document. To read characters with more accuracy high-end OCRs are required which are very costly and may not necessarily be perfectly accurate. There are multiple commonly used OCRs with varying efficiencies; the table 1 given below shows the number of errors they make while recognizing [5]. This is so because the paper document may contain spaces and quotes which are sometimes not recognized by the OCR.

Table 1 Number of errors occurred in commonly used OCRs for a given number of words

		Errors Occurred		
Document	Number of words	Asprise OCR v.4	Adobe Acrobat OCR	Microsoft One Note OCR
1	246	20	0	2
2	140	13	2	4
3	227	18	0	9
4	145	16	1	5
5	235	15	4	10

This paper presents a process that can create a digital signature on paper without using an OCR. This is achieved by using image processing techniques. An image consists of pixels and each pixel of a black and white image can be represented as a one or zero. The above mentioned method is very cost-effective since it has minimum requirements and at the same time maintains all the necessary security features.

2. ASSOCIATED TERMINOLOGY

This section briefs on the various methodologies related to the paper such as asymmetric cryptography, cryptographic hash function, DSA/RSA algorithms, QR codes, image processing.

Asymmetric cryptography uses two separate keys – a public encryption key and a private decryption key to perform opposite functions, each the inverse of the other – as contrasted with symmetric cryptography which relies on the same key to perform both the functions. The keys are related mathematically, but the parameters are chosen so that calculating the private from the public key is either impossible or prohibitively infeasible.

A hash function is a one-way function that maps a variable length message into a fixed length hash digest. There may be more than one message with the same hash value. But this hash valuecannot be mapped to its message. This hash value is basically used to determine whether the message has been altered or not. But it is essential that the hash value must be protected. This is done by encrypting it using RSA or DSA. The hash function used was SHA-1 which converts the message to a 160-bit hash digest. A comparison of the different SHA algorithms is shown in Table 2 [6].

The algorithms that are suitable for digital signatures are DSA, ECDSA, and RSA. DSA is the Digital Signature Algorithm and is based on the difficulty of computing discrete logarithms. ECDSA is Elliptic Curve Digital Signature Algorithm in which the sender and the receiver decide on curve parameters previously. RSA, named after its three creators – Rivest, Shamir and Adleman, is one of the first practicable asymmetric algorithms.

Since it is easily implemented, this paper uses the RSA algorithm.

Table 2 Comparison of SHA parameters

	Various Secure Hash Algorithms			
	SHA-1	SHA-256	SHA-384	SHA-512
Message digest size	160	256	384	512
Message size	$< 2^{64}$	$< 2^{64}$	$< 2^{128}$	$< 2^{128}$
Number of steps	80	64	80	80
Security	80	128	192	256

Notes: 1. All sizes are measured in bits

2.Security refers to the fact that a birthday attackon a message digest of size n produces a collision with a work-factor of approximately $2^{n/2}$

A user of RSA creates and then publishes the product of two large prime numbers, along with an auxiliary value, as their public key. The prime factors must be kept secret. Anyone can use the public key to encrypt a message, but with currently published methods, if the public key is large enough, only someone with knowledge of the prime factors can feasibly decode the message [7]. The digital signature is produced by encrypting the hash digest using RSA as the encryption scheme.The security depends on the difficulty in factoring large numbers. Encryption and decryption are done using equations (1) and (2) where M is the message and C is its corresponding cipher text.

$$C \equiv M^e (mod\ n) \qquad (1)$$

$$M \equiv C^d (mod\ n) = M^{ed} (mod\ n) \qquad (2)$$

Both the sender and the receiver must know the value of n. Thus this is a public key algorithm with public key (e, n) and private key(d, n). [8]

In digital imaging a pixel is the smallest controllable element of a picture on the screen. A black and white image comprises of black and white pixels. The black pixel is represented as a zero and a white pixelas a one, so the image can be represented as a matrix of just ones and zeros. This matrix is converted to a binary string consisting of zeroes and ones. The binary string is subjected to the hash function followed by the encryption.

A QR (Quick Response) code is a 2D barcode which uses four standardized encoding modes (numeric, alphanumeric, byte/binary, and kanji) to efficiently store data. A QR code consists of black square dots arranged on a square grid on a white background, which can be read by an imaging device. The QR code contains information in the vertical as well horizontal direction. QR codes use the Reed-Solomon error correction which can detect and correct multiple errors. QR codes are read using QR scanners and also mobile phones with cameras which have the QR reader software.

3. THE PROPOSED METHOD

In this section, a method to ensure content integrity and authentication of paper documents is presented.

A. Generation process

In this process, the digital signature is produced, converted to a QR code and it is printed on the paper. For generation of the digital signature, the image of certain regions of interest in the document is taken into account. This image is converted to a matrix of ones and zeroes which correspond to the black and white pixels respectively. This data is converted into a binary string which is hashed using a cryptographic hash function like SHA-1 [9]. This hash digest is encrypted using the private key of the signer to create the digital signature of the message [10]. This digital signature is incorporated into a QR code. The QR code is printed on the document and sent to the receiver. This is the digital signature generation process. When the document is being generated, it could preferably have some marked points of reference printed on it which will help the receiver in aligning the paper while scanning. Figure 1 shows the flow chart of the Generation of Digital Signature.

B. Verification process

In this process, the QR code is scanned from the received document to retrieve the digital signature. This signature is decrypted using the public key of the sender to produce a digest. The image of the document is then converted to a binary string which is hashed to produce a digest. This digest is then compared with the decrypted digital signature. If the two digest values match then it can be confirmed that the message content has not been altered and the document is authenticate. Figure 2 shows the flow chart of the Verification of Digital Signature.

In this method, the binary data from the image of the document is used unlike the conventional digital signatures which make use of the content of the document. This ensures that we avoid the shortcomings of character recognition form paper. This method ensures content integrity, that is, in the verification stage one can be sure that the message he has received has not been altered and authenticates the identity of the generator. Only the signer can sign the document since only he knows his private key. The signature can be verified using the corresponding public key which is available to the verifier like the conventional implementation of digital signatures. If the signature is invalid, the verifier cannot authenticate the

message. This could be either because the message was altered or the signer's identity is false.

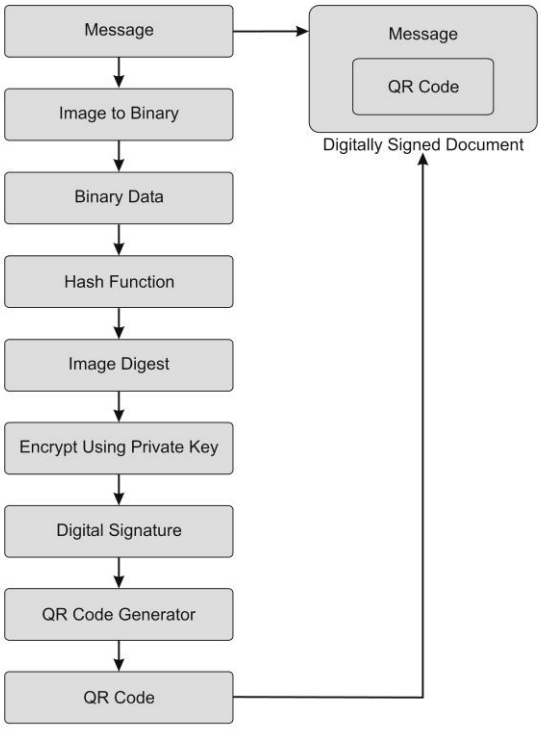

Figure 1 Generation of Digital Signature

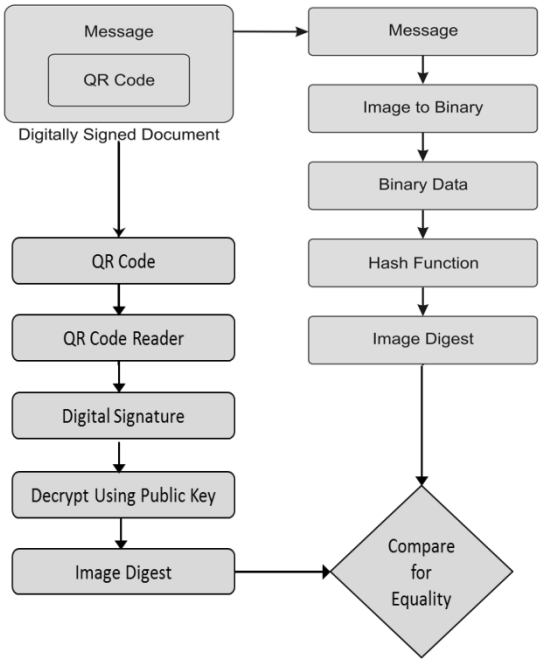

Figure 2 Verification of Digital Signature

If the two hash values obtained in the verification process are not the same, then the message may have been altered. It is infeasible to find a message that produces the same hash value. However this hash value must be protected since an attacker may copy the message and regenerate a hash value. Thus once the message is protected using a private key it is infeasible for an attacker to modify the message and the signature in such a way that it is valid without the knowledge of the sender's private key. This way the integrity is maintained. The signer under no circumstances can deny that he did not send the message because it is with his private key that the message has been encrypted. Only with a corresponding public key can the message be decrypted.

4. IMPLEMENTATION AND DISCUSSION

A system to implement the above mentioned process has been developed using MATLAB and Python [11]. The signer scans the document and selects a region of interest which is easily identified by marking some points of reference on the scanned image. This region of interest is converted to a binary matrix which is in turn converted to a binary string. This is implemented using MATLAB functions. Then the string is hashed using a secure hash algorithm which is SHA-1. The encryption algorithm, RSA, is executed using hashlib functions available with Python. The public and the private keys used in the RSA algorithm can be easily generated by writing a corresponding program in Python.

After this, the digital signature is stored in a QR code. This was accomplished using QR code generator software that was available on the internet. The verifier checks that the document has been unaltered. The document consists of two parts, namely the message and the QR code. The QR code is easily decoded using a mobile phone or a QR scanner. Thus the digital signature is obtained. The signature is then decrypted through RSA algorithm using the signer's public key. The region of interest can be identified from the reference points as marked by the signer. The image of the region of interest is converted to its binary form using MATLAB and the binary string is hashed using the same hash function which is SHA-1. This digest is compared with the hash digest obtained after decrypting the digital signature. If the two match then one can be sure that no alteration has been made to the message.

If a document is expected to be subjected to stamps and other distortions the region of interest should be selected carefully. Only the valued part of the document is considered as the region of interest. Unlike other paper document authentication techniques this method provides for content integrity, and is cost-effective unlike the high-end OCR technology. Also this method provides a way to authenticate handwritten documents, since the document is taken as an image rather than having to convert the handwritten text to digital text using ICR technology.

The physical condition of the paper document should remain intact when received by the verifier because a change in

the document leads to a change in the number of black and white pixels and hence the digital signature. This method does not provide for the originality of the document since it can be easily photocopied as the method deals with only ensuring content integrity. The resolution of the image must remain the same when the document is being verified, because a change in resolution implies a change in the number of pixels and hence different hashes digest. The digital signature generation process is lengthy because it requires two steps, one for scanning the document and another one for printing the reference points and the QR code.

5. CONCLUSION

Content integrity of a paper document is ensured using on-paper digital signatures. A valued portion of the document is selected as a region of interest whose image is converted to binary data and is made use of to generate the digital signature. The method verifies the digital signature efficiently in a cost effective manner. It is also flexible to employ this method on handwritten documents. Any discrepancies that could arise due to misalignment of the paper document while scanning is taken care of by the marked reference points on the document. These help the verifier in identifying the region of interest. Since the digital signature generation process requires two steps, scanning and printing, the attempt to complete the process in one step is still ongoing.

REFERENCE

1. Prashant Kumar Koshta and Dr. Shailendra Singh Thakur, "A novel authenticity of an image using visual cryptography" in International Journal of Computer Science and Network, April 2012.
2. Baoshi Zhu, Jiankang Wu and Mohan S. Kankanhalli, "Print Signatures for Document Authentication".
3. "Print Security", Wikipedia, the free encyclopedia.
4. Ranjan Bose, Information Theory, Coding and Cryptography, 2nd edition, New Delhi, Tata McGraw Hill Education Private Limited, 2008.
5. Maykin Warasart and Pramote Kuacharoen, "Paper-based Document Authentication using Digital Signature and QR Code" in 2012 4th International Conference on Computer Engineering and Technology (ICCET 2012).
6. William Stallings, Cryptography and Network Security Principles and Practices, Prentice Hall, 2005.
7. R.L. Rivest, A.Shamir and L.Adleman, "A method for obtaining Digital Signatures and Public Key Cryptography".
8. "RSA Cryptography Standard", PKCS #1, v2.1, 2002.
9. "Secure Hash Standard", FIPS 180-3, 2008.
10. "Digital Signature Standard" (DSS), FIPS PUB 186-3, 2009.
11. Al Sweigart, Hacking Secret Ciphers with Python, 1st Edition, 2013.